Reconstitution of *Pyrobaculum aerophilum* RNase P, a catalytic ribonucleoprotein involved in tRNA maturation

Honors Research Thesis

Presented in partial fulfillment of the requirements for graduation
with honors research distinction in Biochemistry in the undergraduate colleges of
The Ohio State University

by
Derek Smith

The Ohio State University
May 2011

Project Advisor: Professor Venkat Gopalan, Department of Biochemistry

ABSTRACT

RNase P, an essential ribonucleoprotein (RNP) complex, is responsible for the processing and removal of 5′ tRNA leaders. While RNase P-catalyzed tRNA maturation is common to the three domains of life, the subunit make-up of this enzyme is quite different. The RNase P holoenzyme is composed of one RNase P RNA (RPR) and a variable number of RNase P Protein (RPP) subunits depending on the source: one in bacteria, up to five in archaea, and up to ten in eukarya. While the structurally related RPR is the catalytic moiety in all these RNPs, none of the RPPs (all of which are essential *in vivo*) can catalyze precursor tRNA processing independently of the RPR. Thus, RNase P is a paradigm for the intimate cooperation of RNA and protein subunits necessary to generate an efficient biological catalyst. The failure to reconstitute eukaryotic RNase P (one RPR + ≤ 10 RPPs) has led to the use of archaeal RNase P (one RPR + ≤ 5 RPPs, which share eukaryotic homologs) as an experimental surrogate for studying the functional interplay among multiple protein subunits and a single RNA catalyst. Recent work has demonstrated that *Pyrobaculum aerophilum* (*Pae*), an ancient hyper-thermophilic crenarchaeon, might yet have the smallest natural RPR (207 nucleotides) and only four of the five expected archaeal RPPs.

In this work, the genes of four *Pae* RPPs (POP5, RPP29, RPP30, and L7Ae) were cloned from *Pae* genomic DNA. The first 20 codons of POP5, RPP29, and RPP30 were optimized for protein overexpression in *Escherichia coli* (*Eco*) since the native *Pae* archaeal sequences did not support robust expression in bacteria. POP5 and RPP30 were co-overexpressed as a binary complex in *Eco* SHuffle cells. RPP29 and L7Ae were overexpressed in *Eco* SHuffle and *Eco* BL21(DE3) Rosetta cells, respectively. Successful purification schemes were developed for all four *Pae* RPPs. Ultimately, functional reconstitution of the *Pae* RNase P holoenzyme using *in vitro* transcribed RPR and purified RPPs was achieved.

Overall, *Pae* RNase P is a distinctive variant among its counterparts and this study provides a foundation for dissecting its structure-function relationships and gaining insights into the possible transfer of some functions from the RNA to protein subunits during the evolution of RNase P.

DEDICATION

Dedicated to my parents for their unwavering support and encouragement

ACKNOWLEDGEMENTS

I owe my deepest gratitude to Dr. Venkat Gopalan for his unmatched guidance and his willingness to share a vast knowledge of biochemistry and the scientific process. He has instilled in me the meaning of scientific integrity, the value of being self-critical, and the importance of methodical experimentation.

I also thank Dr. Lien Lai for sharing her expertise through valuable feedback on my work and answers to hundreds of questions. I owe a debt of gratitude to my friends and co-workers in the Gopalan laboratory. I thank Dr. Anil Challa, Dr. Wen-Yi Chen, Dr. I-Ming Cho, Dr. Gireesha Mohannath, Cecilia Go, Stella Lai, and Sathyianarayanan Manivannan for their advice, demonstration of experimental techniques, and most of all for their friendship. My fellow undergraduates, Chigozirim Ekeke and Emily Wong, deserve sincere thanks for their conversations and support that were unlimited even during late hours of the night. I would also like to thank Emily Wong for her efforts in screening and sequencing clones of the *Pae* RPP29 and RPP30 genes.

I am also grateful to Drs. Mark Foster and Amanda Simcox for serving on my honors thesis oral examination committee. I appreciate the generous support of The Ohio State University College of Arts and Sciences for an Undergraduate Research Scholarship (2010-2011), the National Science Foundation: Research Experience for Undergraduates program (2010), and OSU College of Biological Sciences for the Dean's Undergraduate Research Award (2009 and 2010).

In addition, I thank my family and friends for their encouragement throughout my undergraduate career. I especially thank my parents for their steadfast commitment to my education. Their constant support and encouragement have carried me through the most challenging obstacles in life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

°C	Degree Celsius
C domain	Catalytic domain
cm	Centimeter
dCAPS	Derived cleaved amplified polymorphic sequences
ddH_2O	Double distilled water
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
Eco	*Escherichia coli*
EDTA	Ethylenediaminetetracetic acid
FPLC	Fast protein liquid chromatography
GTP	Guanosine triphosphate
h	Hour(s)
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	Isopropyl-β-D-thiogalactopyranoside
kDa	Kilodalton
LB	Luria broth
M	Molar
$MgCl_2$	Magnesium chloride
min	Minute(s)
Mja	*Methanocaldococcus jannaschii*
ml	Milliliter
mM	Millimolar
mRNA	Messenger RNA
Mth	*Methanothermobacter thermauotrophicus*
MWCO	Molecular weight cut off
NaCl	Sodium chloride
Neq	*Nanoarchaeon equitans*
NH_4OAc	Ammonium acetate
nm	Nanometer

nM	Nanomolar
nts	Nucleotides
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
Pae	*Pyrobaculum aerophilum*
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PES	Polyethersulfone
Pfu	*Pyrococcus furiosus*
Pho	*Pyrococcus horikoshii*
PMSF	Phenylmethanesulfonyl fluoride
pre-tRNA	Precursor tRNA
RNA	Ribonucleic acid
RNase P	Ribonuclease P
RNP	Ribonucleoprotein
RPP	RNase P Protein
RPR	RNase P RNA
s	Second(s)
S domain	Specificity domain
SDS	Sodium dodecyl sulfate
Tris	2-Amino-2-hydroxymethyl-propane-1,3-diol
tRNA	Transfer RNA
V	Volt(s)
w/v	Weight per volume
µg	Microgram(s)
µl	Microliter(s)
µm	Micrometer(s)
µM	Micromolar

CHAPTER 1

INTRODUCTION

Highly efficient protein catalysts perform the vast majority of cellular reactions, yet a handful of the most vital functions are the responsibility of ribonucleoproteins (RNPs) that employ RNA catalysts. If we are to understand why these RNA-driven RNPs are retained in a protein-dominated cell, it is critical to elucidate the functional interplay between RNA and protein subunits in these RNPs. Toward this goal, this work was undertaken to attempt functional reconstitution and biochemical characterization of *Pyrobaculum aerophilum* (*Pae*) RNase P, a catalytic RNP. This thesis on the *Pae* RNase P holoenzyme, a distinctive structural variant among its counterparts, describes successful efforts to reconstitute *Pae* RNase P thus setting the stage for an in-depth investigation of its structure-function relationships.

1.1 RNase P

1.1.1 RNase P structure and function

Translation of proteins in the cell relies on tRNAs whose responsibility is the recognition of mRNA codons and delivery of amino acids to the ribosome. Investigation into tRNA biosynthesis in bacteria revealed that tRNAs are made as precursors with extra nucleotide sequences at their 5' ends, which must be removed as part of the tRNA maturation process (1). This gave rise to the question: how is pre-tRNA 5'-processing carried out and what cellular machinery is involved? Nearly four decades ago, it was discovered that a Mg^{2+}-dependent endonuclease was responsible for this cleavage [Fig. 1.1; (2-5)]; this catalytic activity was termed ribonuclease P (RNase P). Subsequent studies identified an RNA moiety as the catalyst in

Escherichia coli (*Eco*) and *Bacillus subtilis* RNase P, and that these RNAs could catalyze precursor-tRNA (pre-tRNA) processing *in vitro* under non-physiological conditions (6). This RNA enzyme in bacterial RNase P, however, requires a protein cofactor for *in vivo* function (7,8). Due to its relatively simple composition (one RNA and one protein) and being the first true natural ribozyme to be uncovered (6), bacterial RNase P has been the focus of extensive biochemical and structural characterization (9-16). Salient insights into the roles of the bacterial RNase P subunits are discussed below.

The bacterial RNase P RNA (RPR) is composed of two independently folding modules: the catalytic (C) and specificity (S) domains, responsible for pre-tRNA cleavage and recognition, respectively (17,18). Long-distance interactions between these two RPR domains help stabilize the tertiary structure (19), and promote concomitant contacts between (i) the pre-tRNA's T stem-loop with the S domain (20), and (ii) the pre-tRNA's 5'-cleavage site with the C domain (21). These molecular struts in the bacterial RPR's tertiary structure juxtapose different parts of the bacterial RPR and facilitate efficient catalysis (11).

As mentioned above, bacterial RNase P employs a single RNase P protein (RPP) cofactor. This bacterial RPP improves substrate recognition by binding the 5' leader of pre-tRNAs, increasing RPR affinity for Mg^{2+} and enhancing the cleavage rate in a substrate identity-dependent fashion (12-14). These important functional alterations supported by the RPP permits RPR function under physiological conditions.

Bacterial RNase P belies the remarkable variability in the subunit composition in all three domains of life. Although the primary function of RNase P is believed to be the same in all the three domains of life, there is a dramatic shift in the make-up of RNase P. Archaeal and eukaryal RNase P (each with one RPR) are associated with up to five and ten RPPs, respectively. The

protein:RNA molecular mass ratio increases from 10% in bacteria to 50% and 70% in archaea and eukarya, respectively, reflecting the increased functional dependence of archaeal/eukaryal RNase P on protein subunits (4, 5, 22). To dissect the interplay between these multiple protein subunits and the RPR, it is vital to reconstitute archaeal/eukaryal RNase P. The latter has proven refractory to date; this has led to the use of archaeal RNase P as an experimental surrogate, especially due to the homology between the five archaeal RPPs and their eukaryal counterparts (23).

The archaeal RNase P holoenzyme is composed of one RPR and up to five RPPs (POP5, RPP21, RPP29, RPP30, and L7Ae) (24). There are three primary classifications for the archaeal RPR: types A, M, and T [Fig. 1.2, (25)]. One significant structural difference between bacterial and archaeal RPRs is the absence in the latter of molecular braces that bridge the two domains in bacterial RPRs (4, 22, 26). It is conceivable that some of the archaeal RPPs play this crucial role supported by RNA-RNA contacts in bacterial RPRs (22, 27).

In previous studies, functional reconstitution of several archaeal RNase P holoenzymes using the five purified RPPs and recombinant RPR was achieved (27-29). Previous research from the Gopalan laboratory has demonstrated that four euryarchaeal RPPs generally function as two binary complexes that interact with the distinct domains of the RPR: POP5•RPP30 with the C domain and RPP21•RPP29 with the S domain (29, 30). A fifth RPP, ribosomal L7Ae, is predicted to bind two kink-turn RNA motifs, one each in the C and S domains (31, 32); this premise remains to be proven. While all five RPPs aid the RPR in pre-tRNA processing, they fulfill distinct roles in the holoenzyme. POP5•RPP30 enhances the rate of pre-tRNA 5'-processing nearly 100-fold while RPP21•RPP29 increases the affinity of the RPR for pre-tRNA substrate by 16-fold (33). Additionally, L7Ae has been shown to increase the thermal

stability of *in vitro* reconstituted *Pyrococcus horikoshii* (*Pho*) RNase P (31). Furthermore, a recent study of a type M RNase P representative has demonstrated that addition of L7Ae to the RPR + POP5•RPP30 + RPP21•RPP29 results in an increase in k_{cat} (turnover number) and decrease in K_m which collectively cause a 360-fold increase in catalytic efficiency (k_{cat}/K_m) (31).

Although high-resolution structural studies have determined the tertiary structures of these five archaeal RPPs, either alone or as binary complexes (29, 34-39), we still lack a structural understanding of the holoenzyme either partly or fully assembled.

1.1.2 *Pyrobaculum aerophilum*: an argument against RNase P ubiquity?

Computational searches based on sequence conservation were utilized by Li and Altman (2004) to uncover over 50 previously unidentified archaeal RNase P RNA sequences but none for *Pyrobaculum* species or *Nanoarchaeon equitans (Neq)*, a parasitic archaeon. Consistent with this finding, Randau et al. (2008) reported that there was no detectable RNase P activity in cell extracts of *Pae* and *Neq*. In both instances, the absence of RNase P in these organisms was attributed to the presence of tRNAs with no 5' leaders, a prediction made from computational studies that examined promoter placement and transcription start sites in tRNA genes. These findings called into question the ubiquity of RNase P and even the long-held belief that all tRNAs are transcribed with additional 5' leader sequences that are subject to RNase P processing. However, the claim that *Pyrobaculum* lacks RNase P has been laid to rest by recent research that revealed a novel variant initially intractable to computational identification.

1.1.3 Identification of *Pyrobaculum aerophilum* RNase P

Dr. Todd Lowe, University of California Santa Cruz, and co-workers identified an RNA that might be the RNA subunit of RNase P in several *Pyrobaculum* species including *Pae* (40). These studies indicated that *Pae* might yet have the smallest natural RPR at 207 nucleotides. This abbreviated RPR has lost most of its S domain yet retains the typical C domain, accounting for its relatively small size compared to other archaeal RPRs (Fig. 1.2). Work in the Gopalan laboratory indicated that indeed this RNA is catalytically active and that a canonical RNase P activity (mediated by an RNP) exists in *Pae* (40). Additional database mining efforts revealed that the RNase P holoenzyme of this ancient hyper-thermophilic crenarchaeon might deviate from the typical euryarchaeal consensus (RPR+5 RPPs). Only four of the five expected archaeal RPPs are computationally identifiable in *Pae* [POP5, RPP29, RPP30, and L7Ae; (40)]. RPP21 seems to be missing - of course, it is conceivable that *Pae* RPP21 has evolutionarily diverged to escape detection by homology searches. It has been shown that archaeal RPP21•RPP29 functions as a binary complex that associates with the S domain of the cognate RPR. These interactions help rationalize the absence of RPP21 in the diminutive *Pae* RPR, which has a minimized S domain. However, whether the absence of RPP21 is a reflection of co-evolution that accompanied the substantially smaller S domain in *Pae* RPR remains unclear.

1.2 Rationale and specific aims

Pae, a unique variant of RNase P, provides a minimal model (Table 1) for uncovering the nuances in structure/composition of this ancient enzyme and merits further investigation due to at least two reasons. First, the *Pae* RNase P holoenzyme catalyzes the same reaction as larger forms of archaeal RNase P. Therefore, it might represent a useful reductionist model to understand the function of the individual RNase P subunits. It may even be better suited for

high-resolution structural studies. Second, *Pae* RNase P might yet provide new insights into the dynamic co-evolution of enzymes and substrates, since the smaller size of *Pae* RNase P could be an adaptation that is well suited for a specific suite of pre-tRNAs (perhaps with unusually short leaders as in *Pyrobaculum*) or other non-coding RNAs (40). The overall goal of this study is functional reconstitution and biochemical characterization of the *Pae* RNase P holoenzyme. The three specific aims are: (i) cloning of the four *Pae* RPP-encoding genes, (ii) overexpression and purification of the four *Pae* RPPs, and (iii) functional reconstitution of *Pae* RNase P.

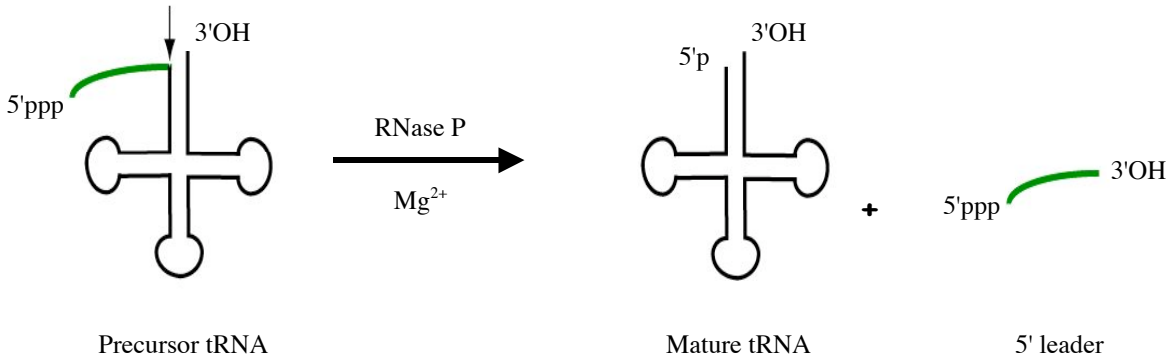
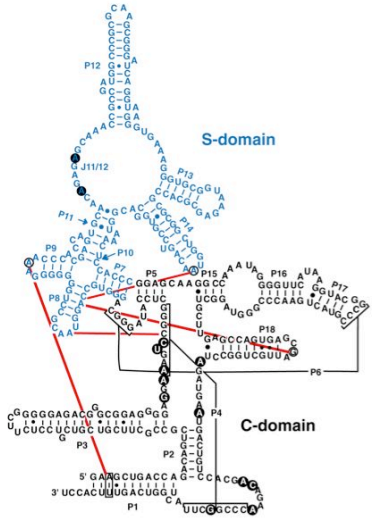
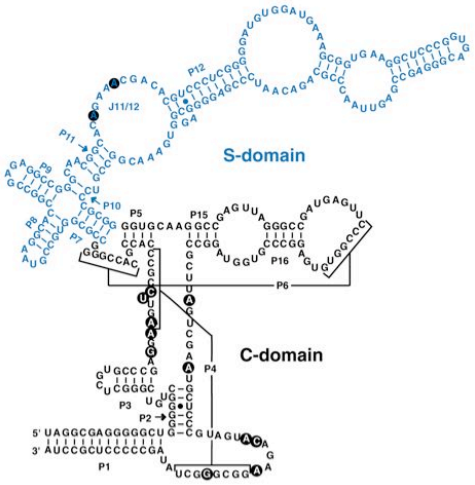


Figure 1.1. 5' Maturation of tRNAs by RNase P.

RNase P-catalyzed, Mg^{2+}-dependent removal of 5' leaders from pre-tRNAs to yield mature tRNAs. Figure modified from ref. (22).

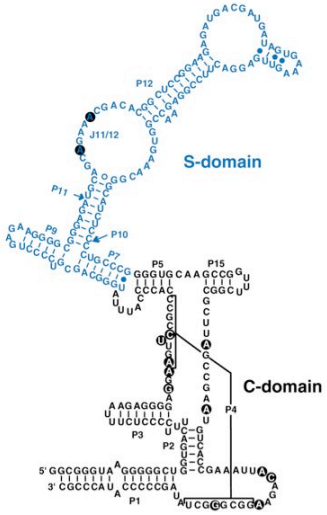


Escherichia coli (*Eco*) RPR – 377 nts



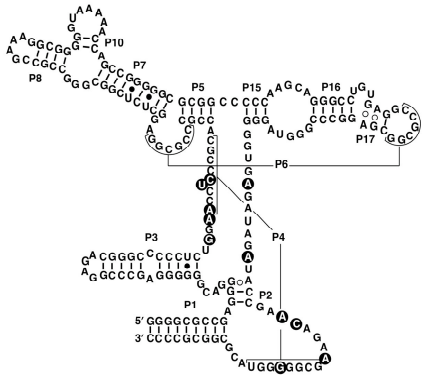
Pyrococcus furiosus (*Pfu*) RPR – 330 nts

Type A



Methanocaldococcus jannaschii (*Mja*) RPR – 267 nts

Type M



Pyrobaculum aerophilum (*Pae*) RPR – 207 nts

Type T

Figure 1.2. Structural comparison of bacterial and archaeal RPRs.

Comparison of the bacterial *Eco* RPR versus archaeal *Pfu* and *Mja* RPRs highlights the common two-domain structure of RNase P RNA subunits. In addition, note the substantial reduction of the S domain in archaeal type T RPRs as compared to archaeal type A and M RPRs. Figure modified from ref. (41).

Model system	Subunit	Molecular mass (kDa)	Isoelectric point (pI)
Pae	RPR	68	-
	POP5	9	10.1
	RPP21	-	-
	RPP29	9	8.8
	RPP30	20	5
	L7Ae	16	8.9
Pfu	RPR	110	-
	POP5	14	10.7
	RPP21	14	11
	RPP29	15	10.8
	RPP30	25	10.3
	L7Ae	14	5.1
Mth	RPR	100	-
	POP5	15	10.5
	RPP21	17	11.1
	RPP29	11	11.4
	RPP30	28	8.2
	L7Ae	14	4.4
Mja	RPR	87	-
	POP5	16	10
	RPP21	16	10.6
	RPP29	11	9.7
	RPP30	27	9.1
	L7Ae	13	5.5

Holoenzyme	Molecular mass (kDa)
Pae	122
Pfu	191
Mth	184
Mja	169

Table 1. Comparison of the molecular weight and isoelectric point of RNase P subunits from different archaea.

Alignments of the *Pae* RPR and RPP sequences with other archaeal counterparts are provided in ref. 40.

CHAPTER 2

2.1 Materials and methods

2.1.1 Cloning the genes encoding the *Pae* RPPs

The genes encoding *Pae* POP5, RPP29, RPP30, and L7Ae were amplified by PCR using *Pae* genomic DNA as the template and the respective gene-specific primers listed in Table 2. All PCR products were purified and digested with *Bam*HI (whose recognition site was incorporated in the reverse primers). The four *Bam*HI-treated PCR products were processed differently depending on the ORF being cloned: digestion with *Nco*I (POP5) or *Nde*I (RPP29, RPP30, and L7Ae). The POP5-encoding ORF was ligated to pET-15b (Novagen) that had been digested with *Nco*I and *Bam*HI. The RPP29- and RPP30-encoding ORFs were separately ligated to pLANT2b (42) that had been digested with *Nde*I and *Bam*HI. The L7Ae-encoding ORF was ligated to pET-33b (Novagen) that had been digested with *Nde*I and *Bam*HI. A small aliquot from each ligation reaction was used to transform electrocompetent *Eco* DH5α cells. Transformants were screened for candidate clones and positives were identified using automated DNA sequencing at the Ohio State University Plant Microbe Genomics Facility. The screening and sequencing of *Pae* RPP29 and RPP30 clones was completed by Emily Wong, an undergraduate in the Gopalan laboratory (summer 2009).

2.1.2 Codon optimization of *Pae* RPP genes for *Eco* overexpression

Initial attempts to overexpress *Pae* POP5, RPP29, and RPP30 in *Eco* cells were unsuccessful likely due to differences in codon usage between *Pae* and *Eco* genes. Highly expressed genes during the exponential growth phase of *Eco* (class II genes) display a codon bias towards high-frequency tRNA isoacceptors (Novagen: Codon Usage Table). Therefore, codon usage in

native sequences of *Pae* POP5, RPP29, and RPP30 genes were compared against the codon usage in *Eco* class II genes using the Graphical Codon Usage Analyzer (http://gcua.schoedl.de/). The first 20 codons of each *Pae* RPP gene with an *Eco* class II frequency ≤ 40% were optimized to the highest frequency *Eco* class II codon with respect to the initial amino acid sequence [excluding optimization of TCT (32%) and GCG (32%) which are highest frequency for their respective amino acids, and GAG (25%) due to certain primer design requirements].

Codon optimization was performed by site-directed mutagenesis with two gene-specific primers designed to incorporate nucleotide mismatches at sites of desired codon replacement. Codon-optimized genes were generated by touchdown PCR (whole-plasmid amplification) using the native copy (section 2.1.1) as the template and the respective phosphorylated gene-specific primers listed in Table 2. All primers were phosphorylated with Optikinase (USB) for 30 min at 37 °C. Briefly, the method employs two phosphorylated primers designed to abut each other on the template but face outward. These primers were used to incorporate the desired changes in each RPP-encoding gene during PCR amplification of the full plasmid. These PCR products were circularized and used to transform *Eco* DH5α cells. *Pae* POP5 and RPP29 transformants were screened for candidate clones using unique restriction sites created as a result of codon optimization. *Pae* RPP30 did not create any unique restriction site as a result of codon optimization; therefore, a unique restriction site for *Bsa*I was created by dCAPS (Derived Cleaved Amplified Polymorphic Sequences) to facilitate screening (adapted from ref. 43). This PCR-based technique introduces a single nucleotide polymorphism within a gene using two sequence-specific primers with one primer containing a single nucleotide mismatch at the desired site of mutation. Small aliquots of isolated DNA from *Pae* RPP30 transformants were amplified using dCAPS primers and screened using *Bsa*I. A positive clone for the optimized *Pae* RPP30

gene was subsequently identified. Putative positive clones for all three optimized *Pae* RPP genes were validated using automated DNA sequencing at the Ohio State University Plant Microbe Genomics Facility.

2.1.3 Overexpression of *Pae* RPPs

Pae POP5 and RPP30 were co-overexpressed from codon-optimized sequences to generate a binary complex (POP5•RPP30) *in vivo*. Both RPP29 and L7Ae were overexpressed individually from codon-optimized and native sequences, respectively. *Eco* SHuffle cells* co-transformed with pET-15b-POP5(optimized) and pLANT2b-RPP30(optimized) were used to inoculate 5 ml Luria broth (LB) media containing 100 μg/ml carbenicillin and 35 μg/ml kanamycin. *Eco* SHuffle cells transformed with pLANT2b-RPP29(optimized) and *Eco* BL21(DE3) cells transformed with pET-33b-L7Ae were used to separately inoculate 5 ml LB media containing 35 μg/ml kanamycin. These three cultures were grown 16 h at 37 °C with shaking and used to inoculate 500 ml LB media supplemented with appropriate antibiotics. The cultures were grown at 37 °C with shaking to an OD_{600} of 0.5 - 0.7. Protein overexpression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown for an additional 4 h at 37 °C with shaking. Subsequently, the cells were harvested by centrifugation and stored at - 80 °C.

---

*Rationale for the use of *Eco* SHuffle cells is provided in section 2.2.2.

2.1.4 Purification of *Pae* RPPs

All chromatographic purifications were performed on a ÄKTA FPLC purifier system using 1-ml FPLC columns (ion-exchange or hydrophobic matrices) purchased from GE Healthcare. The individual purification schemes employed for the different RPPs are discussed below (Fig. 2.5).

Pae POP5

Frozen cells were thawed on ice and then re-suspended in 20 ml buffer A [25 mM Tris-HCl (pH 7.5), 1 mM EDTA, 5 mM DTT, 0.1 mM PMSF] supplemented with 7 M urea. The re-suspended cells were lysed by sonication [time: 4 min, program: 4 s on and 10 s off, amplitude: 70% (Sonics, VCX130)] and cell debris pelleted by centrifugation (9,700×g, 15 min, 4 °C). The supernatant was filtered through a 0.45 μm PES filter and concentrated with a 15-ml Centriprep 10-kDa molecular weight cut off (MWCO) centrifugal filter (3,000×g, 200 min, 4 °C). The concentrated proteins were separated by SDS-PAGE and visualized with reversible zinc staining (adapted from ref. 44). The protein band corresponding to POP5 was excised and washed twice with phosphate-buffered saline (pH 7.4) containing 100 mM EDTA. POP5-containing gel fragments were re-suspended in 0.5 ml buffer A supplemented with 4 M urea, enclosed in dialysis tubing (BioDesignDialysis Tubing, D302), and electroeluted against 300 ml re-suspension buffer (25 V, 5.5 h, 22 °C). After pelleting gel fragments by centrifugation, electroeluted POP5 was refolded by gradual removal of urea, which was accomplished by step-down dialysis. POP5 was dialyzed against assay buffer [50 mM HEPES (pH 7.5) and 800 mM NH_4OAc] and stored on ice at 4 °C until further use. The concentration of POP5 was calculated using the predicted extinction coefficient (6,640 cm^{-1} M^{-1}, http://ca.expasy.org) at 280 nm.

Pae RPP30

Frozen cells were thawed on ice and then re-suspended in 20 ml buffer A supplemented with 7 M urea. The re-suspended cells were lysed by sonication (time: 4 min, program: 4 s on and 10 s off, amplitude: 70%) and the cell debris was pelleted by centrifugation (9,700×g, 15 min, 4 °C). The supernatant was filtered through a 0.45 μm PES filter and RPP30 was refolded by

gradual removal of urea, which was accomplished by step-down dialysis. The re-folded protein was loaded to a pre-equilibrated 1-ml HiTrap SP-Sepharose FF column and collected in the flow-through, which was then loaded to a pre-equilibrated 1-ml HiTrap Q-Sepharose FF column. RPP30 was eluted with a 0 to 2 M NaCl gradient in buffer A. The presence of RPP30 in fractions corresponding to 0.5 - 0.9 M NaCl was identified using SDS-PAGE and Coomassie blue staining (Fig. 2.6). These peak fractions, which had some minor contaminants, were pooled and the NaCl concentration increased to 3 M in preparation for hydrophobic chromatography. The pooled fractions were loaded to a pre-equilibrated 1-ml HiTrap phenyl-Sepharose HP column and eluted with a 3 to 0 M reverse NaCl gradient in buffer A. RPP30 eluted between 0.8 - 0.5 M NaCl and was judged to be homogeneous by SDS-PAGE and Coomassie blue staining (Fig. 2.7). Fractions containing purified RPP30 were pooled, dialyzed against assay buffer [50 mM HEPES (pH 7.5) and 800 mM NH_4OAc], and stored on ice at 4 °C until further use. The concentration of RPP30 was calculated using the predicted extinction coefficient (36,130 cm^{-1} M^{-1}) at 280 nm.

Pae RPP29

Pae RPP29 was purified following the protocol described above for *Pae* POP5 with the exception that 10 mM DTT was used in the sonication buffer. Also, RPP29 is predicted to have an extinction coefficient of 6,680 cm^{-1} M^{-1} at 280 nm.

Pae L7Ae

Frozen cells were thawed on ice and then re-suspended in 25 ml buffer A containing 1 M NaCl. The re-suspended cells were lysed by sonication (time: 4 min, program: 4 s on and 10 s off, amplitude: 60%) and the cell debris was pelleted by centrifugation (9,700×g, 15 min, 4 °C). The supernatant was filtered through a 0.45 μm PES filter. The clarified supernatant was treated with

PEI to a final concentration of 0.05% (v/v) and incubated on ice for 30 min (Fig. 2.8). PEI-precipitated nucleic acid was pelleted by centrifugation (9,700×g, 15 min, 4 °C). The PEI supernatant was heat-treated at 75 °C for 20 min (Fig. 2.8). Heat-precipitated *Eco* proteins were pelleted by centrifugation (9,700×g, 15 min, 4 °C). The soluble fraction was dialyzed against buffer A containing 25 mM NaCl, loaded to a pre-equilibrated 1-ml HiTrap SP FF Sepharose column, and eluted with a 0 to 2 M NaCl gradient in buffer A (Fig. 2.8). SDS-PAGE followed by Coomassie blue staining was used to identify the fractions containing L7Ae (Fig. 2.9). Typically L7Ae fractions eluted between 0.9 - 1.3 M NaCl. These fractions were pooled and the NaCl concentration was increased to 2.5 M in preparation for hydrophobic chromatography. The pooled fractions were loaded to a pre-equilibrated 1-ml HiTrap Phenyl Sepharose HP column and eluted with a reverse 2.5 to 0 M NaCl gradient in buffer A. L7Ae eluted at ~2.5 M NaCl (i.e., at the beginning of the reverse gradient) and was judged to be homogeneous by SDS-PAGE and Coomassie blue staining (Fig. 2.10). Fractions containing purified L7Ae were pooled, dialyzed against assay buffer [50 mM HEPES (pH 7.5) and 800 mM NH_4OAc], and stored on ice at 4 °C until further use. The concentration of L7Ae was calculated using the predicted extinction coefficient (5,960 cm^{-1} M^{-1}) at 280 nm.

2.1.5 *In vitro* transcription: *Pae* RPR and pre-tRNA^{Phe}

The gene encoding the *Pae* RPR was amplified by PCR using as template pBT7-*Pae* RPR (40) and as primers M13 forward-ext (forward) and R-ext (reverse; Table 2). The PCR product was digested with *Sfu*I before its use as a template in an *in vitro* T7 RNA polymerase-mediated runoff transcription reaction. The transcription product was treated with RNase-free DNase I for 1 h at 37 °C, phenol:chloroform extracted, and extensively dialyzed against distilled water. The RNA was ethanol precipitated, pelleted by centrifugation (14,000×g, 5 min), air dried, and

re-suspended in 10 μl ddH_2O. The RNA concentration was calculated using the predicted extinction coefficient (1,698,700 cm^{-1} M^{-1}, www.ambion.com) at 260 nm.

The gene encoding *Pae* pre-$tRNA^{Phe}$ was amplified by PCR using pBT7-*Pae* $tRNA^{Phe}$ as template and the primers M13 forward-ext (forward) and *Pae*Phe-2R (reverse; Table 2). The PCR product was used as the template in an *in vitro* transcription reaction, which was processed as described above for the *Pae* RPR. To prepare internally radiolabeled pre-$tRNA^{Phe}$, the *in vitro* transcription reaction was performed in the presence of [α-^{32}P]-GTP and the radiolabeled RNA gel purified and ethanol precipitated.

2.1.6 *Pae* RNase P multiple-turnover assays

RNase P activity was assayed at 55 °C for 3 h using 100 nM *in vitro* transcribed, folded RNA and 1000 nM of each purified RPP in 50 mM HEPES (pH 7.5), 5 mM $MgCl_2$, and 800 mM NH_4OAc. Multiple-turnover assays were conducted with 500 nM *in vitro* transcribed pre-$tRNA^{Phe}$ mixed with a trace amount of radiolabeled pre-$tRNA^{Phe}$. The cleavage products were separated on a 10% (w/v) polyacrylamide/7 M urea gel and radiographed on a Typhoon Trio Variable Mode Phosphorimager (GE Healthcare). Cleavage products were quantitated using ImageQuant software.

2.2 Results and discussion

2.2.1 Comparative analysis of *Pae* and *Eco* codon usage for optimization of genes encoding *Pae* POP5, RPP29, and RPP30

Initial attempts to overexpress *Pae* POP5, RPP29, and RPP30 using their native gene sequences (Fig. 2.1) in *Eco* cells were unsuccessful. To address this limitation, a methodical analysis of codon usage and correlation was undertaken. Protein overexpression in a culture of *Eco* host cells is typically induced during the exponential growth phase (OD_{600} ~ 0.6), a period during which a group of genes termed *Eco* class II genes are highly expressed. The frequency of codon usage within this set of genes has been determined to provide an indication of which codons support high expression. These values were used to analyze *Pae* gene sequences to determine how effectively *Eco* cells could synthesize *Pae* RPPs (Table 3).

It is obvious from this analysis that translation of *Pae* RPP native mRNAs will entail use of several low frequency tRNA isoacceptors. Codon correlation, the frequency of tRNA changes required between codons for the same amino acid, has also been shown to play a significant role in translation (45). Further insight into *Eco* translation of *Pae* RPPs was gleaned from frequency and correlation analysis of the first 20 codons of each *Pae* mRNA transcript. Based on these results, primers complementary to the first 60 nucleotides of each RPP gene (first 20 codons of each mRNA transcript) were designed to optimize codons with an *Eco* class II frequency ≤ 40% [excluding the codons TCT (32%) and GCG (32%) which are the highest frequency codon for their respective amino acids]. These primers were used to mutagenize nine POP5 codons, 14 RPP29 codons, and 14 RPP30 codons (Figs. 2.2 and 2.3). A correlation analysis of the optimized gene sequences revealed a net reduction of three tRNA changes in POP5 and no net change in

RPP29 and RPP30. Improved codon usage and correlation was borne out when all three RPPs were successfully overexpressed from codon-optimized genes (Fig. 2.4).

2.2.2 Overexpression and purification of *Pae* RPPs

Overexpression of *Pae* POP5, RPP29, and RPP30 from codon-optimized gene sequences was initially accomplished in *Eco* BL21(DE3) Rosetta cells; however, after sonication these RPPs remained in the insoluble fraction. It was hypothesized that the reducing environment of the cytoplasm might inhibit proper protein folding by disrupting disulfide bond formation. Analysis of the primary structure of POP5, RPP29, and RPP30 reveals two, four, and zero cysteine residues, respectively. Although RPP30 does not contain any cysteine residues, it was overexpressed as a binary complex with POP5 (contains two cysteine residues) possibly resulting in its observed insolubility. To improve conditions for proper protein folding, all three RPPs were overexpressed in *Eco* SHuffle cells, which offer an oxidizing cytoplasm amenable to disulfide bond formation. Although overexpression of the three *Pae* RPPs in these cells did not fully resolve the issue of insolubility, we did resort to the use of *Eco* SHuffle cells since there were some improvements in protein overexpression potentially due to improved folding.

Although only modest amounts of *Pae* POP5, RPP29, and RPP30 were soluble, we were able to recover these proteins from the inclusion bodies by denaturing these proteins in 7 M urea-containing buffer during sonication. To facilitate proper protein refolding, urea was gradually removed by dialysis using 1 M step-downs of 1 h duration. Ultimately, the four *Pae* RPPs were successfully purified using varied combinations of ion-exchange and hydrophobic chromatography, heat treatment, and PAGE-based isolation (Fig. 2.11). With the four purified protein subunits in hand, the final aim of this work was within reach.

2.2.3 Functional reconstitution of *Pae* RNase P

The final aim and overarching goal of this study was achieved upon successful reconstitution of the *Pae* RNase P holoenzyme (Fig. 2.12). A multiple turnover pre-tRNA processing assay was preformed to examine whether the *Pae* RPR and four purified *Pae* RPPs could function as a holoenzyme *in vitro*. *Pae* RNase P was reconstituted using 100 nM *in vitro* transcribed, folded RPR and 1000 nM of each purified RPP (assay conditions are described in section 2.1.6.). Indeed, 5'-processing of pre-tRNA^{Phe} was observed for reconstituted *Pae* RNase P. Two negative controls were performed to verify that pre-tRNA processing activity requires both the RPR and RPPs: *Pae* RPR alone (Fig. 2.12, lane 2) and the four *Pae* RPPs alone (data not shown). Neither reaction indicated 5'-tRNA processing activity under the conditions used for holoenzyme reconstitution. Additionally, *Eco* RNase P was employed as a positive control for pre-tRNA^{Phe} 5'-processing. The observed *Pae* RNase P activity marks the first functional reconstitution of the smallest, naturally occurring RNase P.

2.3 Future directions and concluding remarks

The cloning, overexpression, and purification of the four *Pae* RPPs were critical steps towards the functional reconstitution of the *Pae* RNase P holoenzyme and have established a framework for characterizing this unique RNase P variant. In addition to this preliminary work, additional measures to ensure the homogenity of the gel-purified POP5 and RPP29 (e.g., two-dimensional gel electrophoresis) would provide definitive proof that the isolated proteins are indeed homogenous.

The conditions for optimal *Pae* RNase P activity will need to be identified by varying the concentration of monovalent and divalent ions, pH, and temperature. To identify which RPPs

play a critical role in chemical cleavage, substrate recognition, and thermostability, the 15 possible RPR±RPP variations must be examined in a combinatorial fashion.

Furthermore, mutagenesis of the *Pae* RPR should be undertaken to investigate potential disruptions in binding interactions with L7Ae. Previous work in the Gopalan laboratory and by others has shown that typical archaeal RNase P RPRs have two L7Ae-binding sites, one each in the C and S domains (31, 32). Interestingly, the *Pae* RPR contains only one of these binding sites (i.e., in the C domain) due to its smaller S domain. PCR-based mutagenesis of the two G•A pairs (in the L7Ae-recognized kink-turn) to G•C pairs would permit testing of L7Ae binding to the *Pae* RPR and add another dimension to the structural understanding of this unique RNase P variant.

Overall, this work has laid a solid foundation for further studies that will dissect the nuances of the relationship between structure/composition and function of RNase P.

1. *Pae*POP5-F	5' – GACTTCATGGGGTGCTACGCGG – 3'
2. *Pae*POP5-R	5' – GATGGATCCTTAGTGTTTATATAAGGCTTTTCTTATCGACG – 3'
3. *Pae*RPP29-F	5' – GTCACATATGCCTTTTAGTTGTATCGACTTGC – 3'
4. *Pae*RPP29-R	5' – GTCAGGATCCTTAGAATTTGCCGCAGTTGAATAG – 3'
5. *Pae*RPP30-F	5' – GTCACATATGGCCTTTGTGATTAAGAGGG – 3'
6. *Pae*RPP30-R	5' – GTCAGGATCCTTAGTCATCTCTGCTCGCCCA – 3'
7. *Pae*L7Ae-F	5' – GATCCATATGGCAGTGACCATCGACCCAAAGACGTTCTACGCA – 3'
8. *Pae*L7Ae-R	5' – GATCGGATCCTTACCTCTTTGCTGGCACAGG – 3'
9. *Pae*POP5opt-F	5' – TACCAGCTGATCGGCTTCCAGGCGCTGACCCGCCCGCCTAG – 3'
10. *Pae*POP5opt-R	5' – ACGTTCCAGGATTTCCGCGTAGCAACCCATGGTATATCTCC – 3'
11. *Pae*RPP29opt-F	5' – CGTGAAGTTTTCACCTACAAATGCCCGTACGACGTTAGGGGG – 3'
12. *Pae*RPP29opt-R	5' – ACCCAGCAGGTCGATGCAAGAGAACGGCATATGTATATCTCCTTC – 3'
13. *Pae*RPP30opt-F	5' – GAATGGGACTTCGCGGTTTACTCTCCGGAGGTAGAAAAGGCG – 3'
14. *Pae*RPP30opt-R	5' – AACGAAACCACGTTTGATACCGAACGCCATATGTATATCTCCTTC – 3'
15. F-RPP30opt-dCAPS	5' – GTTATCAAACGTGGTCTCGTTGAATGG – 3'
16. R-RPP30opt-dCAPS	5' – ATATCAACTCCGCGGACAAAAGGC – 3'
17. M13 forward-ext	5' – CGACGTTGTAAAACGACGGCCAG – 3'
18. R-ext	5' – GGAAACAGCTATGACCATGATTACGCCAAG – 3'
19. *Pae*Phe-2R	5' – TGGTGCGGCCGCCGG – 3'

Table 2. Oligonucleotides used in this study.

Oligonucleotides 1 through 8 were used for cloning of the genes encoding *Pae* RPPs. Oligonucleotides 9 through 14 were used for codon optimization of genes encoding *Pae* POP5, RPP29, and RPP30. Oligonucleotides 15 and 16 were used to incorporate a unique restriction site into RPP30 optimized clones by dCAPS for screening. Oligonucleotides 17 through 19 were used for amplification of *Pae* RPR and pre-tRNA^{Phe} genes, which then served as templates for *in vitro* transcription.

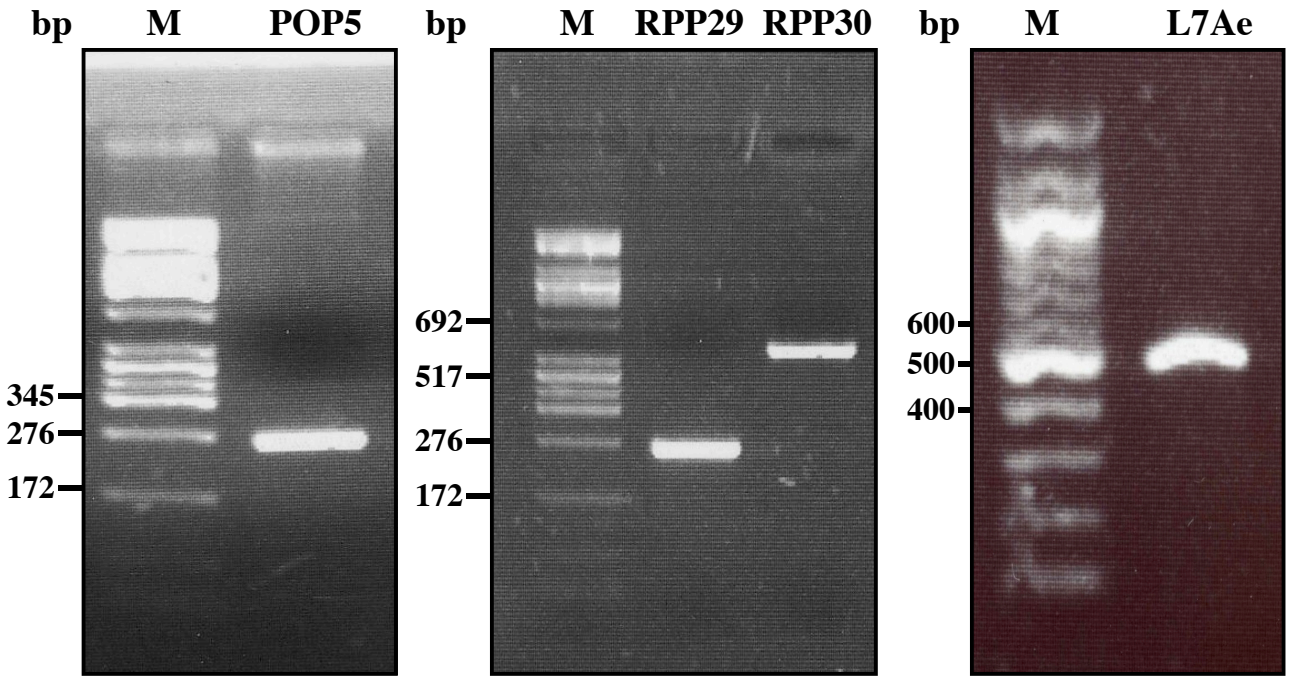


Figure 2.1. Cloning of genes encoding *Pae* RPPs.

PCR-based amplification of *Pae* POP5, RPP29, RPP30, and L7Ae genes using *Pae* genomic DNA as the template. M: molecular size markers/standards

(A)

Gene	Number of codons in the entire transcript with an *Eco* class II frequency ≤ 10%	Fraction of low-frequency codons
Pae POP5	22	27%
Pae RPP29	26	27%
Pae RPP30	43	24%

Gene	Number of codons in the entire transcript with an *Eco* class II frequency ≤ 25%	Fraction of low-frequency codons
Pae POP5	38	46%
Pae RPP29	39	48%
Pae RPP30	88	50%

Gene	Number of codons in the entire transcript with an *Eco* class II frequency ≤ 40%	Fraction of low-frequency codons
Pae POP5	57	70%
Pae RPP29	51	70%
Pae RPP30	128	72%

(B)

Gene	Number of codons within the first 20 of each transcript with an *Eco* class II frequency ≤ 10%	Fraction of low-frequency codons
Pae POP5	6	30%
Pae RPP29	5	25%
Pae RPP30	3	15%

Gene	Number of codons within the first 20 of each transcript with an *Eco* class II frequency ≤ 40%	Fraction of low-frequency codons
Pae POP5	9	45%
Pae RPP29	14	70%
Pae RPP30	14	70%

Table 3. Analysis of codon usage in *Pae* RPP genes.

Analysis of codon usage in *Pae* RPP-encoding genes based on the frequency of codon usage within the *Eco* class II genes. A whole-gene analysis of codon usage is presented in panel (A) while panel (B) lists a local analysis of the first 20 codons of each RPP-encoding gene.

POP5 gene: native and optimized

ATG GGG TGC TAC GCG GAG ATC CTC GAG AGG TAT CAG CTA ATC GGC TTC CAG GCG CTA ACA
ATG GGT TGC TAC GCG GAA ATC CTG GAA CGT TAC CAG CTG ATC GGC TTC CAG GCG CTG ACC

RPP29 gene: native and optimized

ATG CCT TTT AGT TGT ATC GAC TTG CTG GGG AGA GAG GTA TTT ACA TAT AAG TGC CCT TAC
ATG CCG TTC TCT TGC ATC GAC CTG CTG GGT CGT GAA GTT TTC ACC TAC AAA TGC CCG TAC

RPP30 gene: native and optimized

ATG GCC TTT GTG ATT AAG AGG GGG TTT GTG GAG TGG GAC TTC GCC GTG TAT TCT CCC GAG
ATG GCG TTC GTT ATC AAA CGT GGT TTC GTT GAA TGG GAC TTC GCG GTT TAC TCT CCG GAG

Figure 2.2. Analysis of *Pae* POP5, RPP29, and RPP30 native gene sequences for the presence of codons used infrequently in *Eco*.

A comparative analysis of codon usage in *Eco* class II genes and the first 60 nucleotides of *Pae* POP5, RPP29, and RPP30 native genes (equivalent to the first 20 codons of *Pae* mRNA transcripts). All codons highlighted in red represent native codons in *Pae* genes with a usage frequency ≤ 40% in *Eco* class II genes. All codons highlighted in green represent sites optimized by site-directed mutagenesis to improve protein translation efficiency in *Eco* cell cultures.

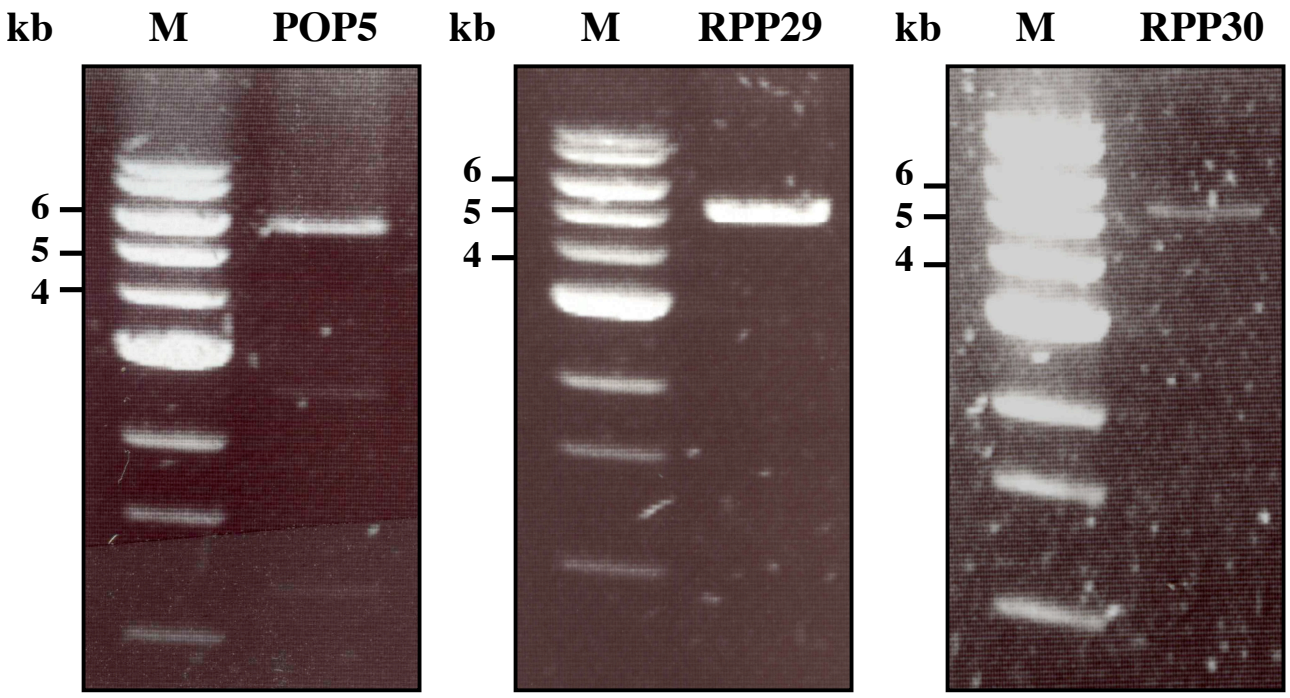


Figure 2.3. Codon optimization of genes encoding *Pae* POP5, RPP29, and RPP30.

Codon optimization of *Pae* POP5, RPP29, and RPP30 encoding genes by whole-plasmid touchdown PCR-based site-directed mutagenesis using as templates the vectors that contain the respective native genes. M: molecular size markers/standards

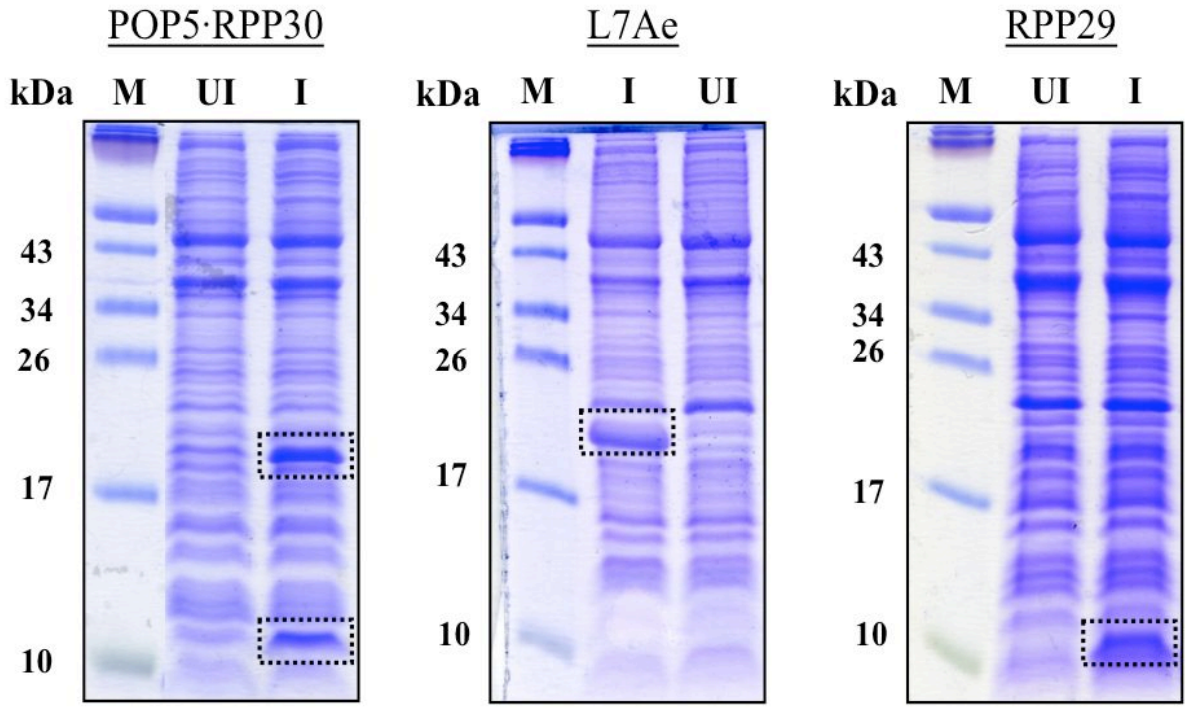


Figure 2.4. Overexpression of *Pae* RPPs.

SDS-PAGE analysis of *Pae* POP5, RPP30, and RPP29 overexpression in *Eco* SHuffle cells from codon-optimized gene sequences; *Pae* L7Ae overexpression was performed in *Eco* BL21(DE3) cells from a native gene sequence. M: molecular weight markers/standards; UI: un-induced; I: induced (1 mM IPTG).

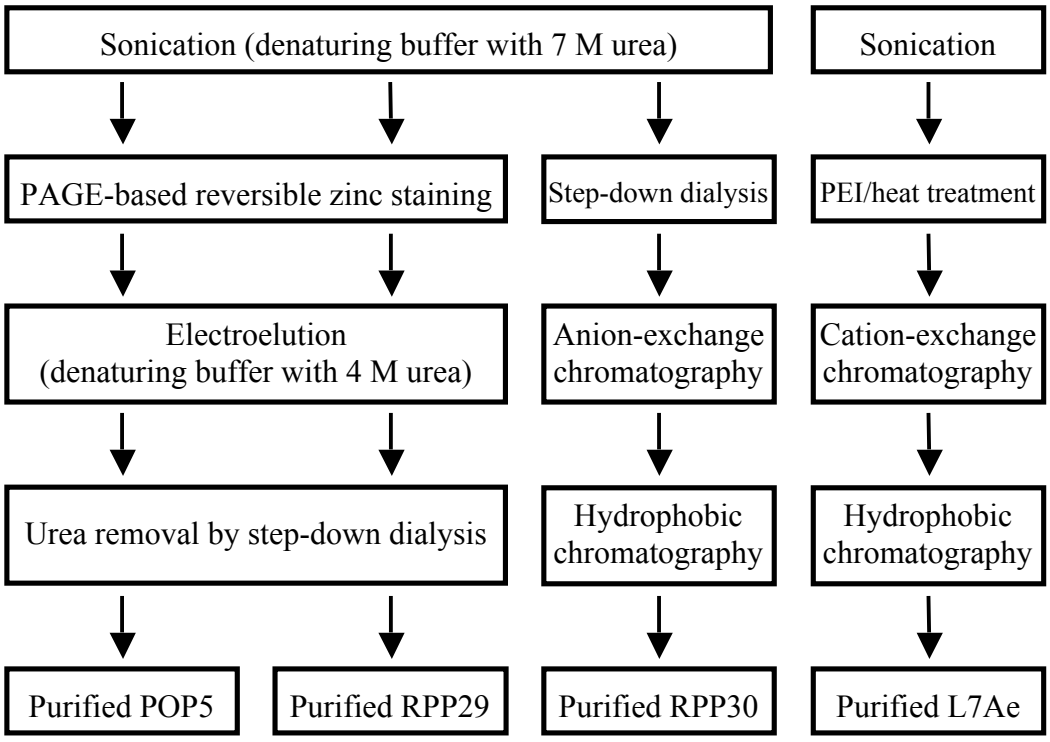


Figure 2.5. Overall schematic depicting the different schemes employed for purification of *Pae* RPPs.

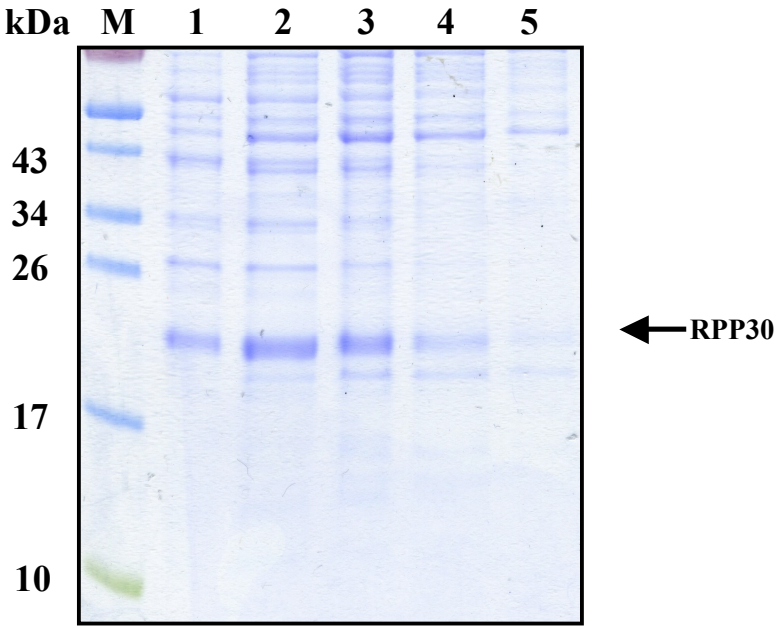


Figure 2.6. SDS-PAGE analysis of *Pae* RPP30 in the anion-exchange chromatography eluent fractions.

Pae RPP30 eluted from a 1-ml HiTrap Q-Sepharose FF column over a 0 to 2 M NaCl gradient. *Pae* RPP30 eluted in the range of approximately 0.5 - 0.9 M NaCl. Fractions containing *Pae* RPP30 were identified using 15% polyacrylamide tricine-SDS-PAGE and Coomassie blue staining. M: molecular weight markers/standards; 1-5: *Pae* RPP30 containing fractions.

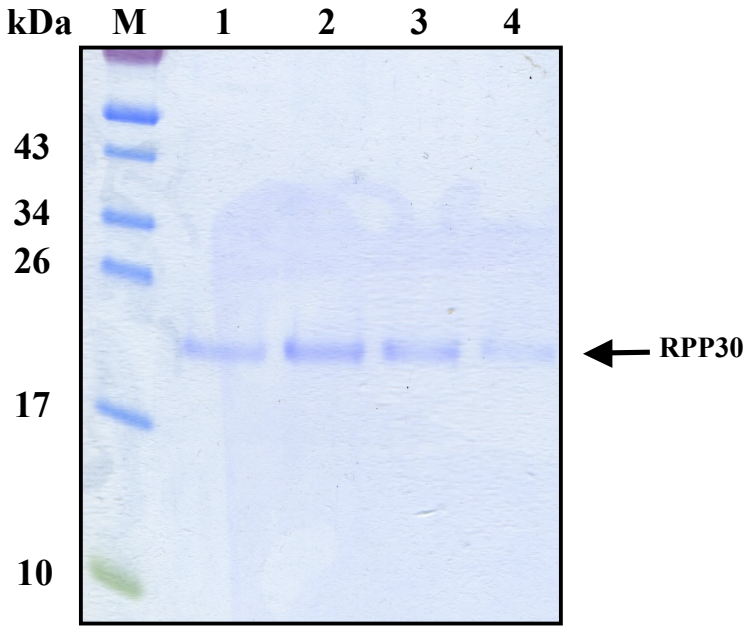


Figure 2.7. SDS-PAGE analysis of *Pae* RPP30 in the hydrophobic chromatography eluent fractions.

Pae RPP30 eluted from a 1-ml HiTrap phenyl-Sepharose HP column over a 3 to 0 M NaCl gradient. *Pae* RPP30 eluted in the range of approximately 0.8 - 0.5 M NaCl. Fractions containing *Pae* RPP30 were identified using 15% polyacrylamide tricine-SDS-PAGE and Coomassie blue staining. M: molecular weight markers/standards; 1-4: *Pae* RPP30 containing fractions.

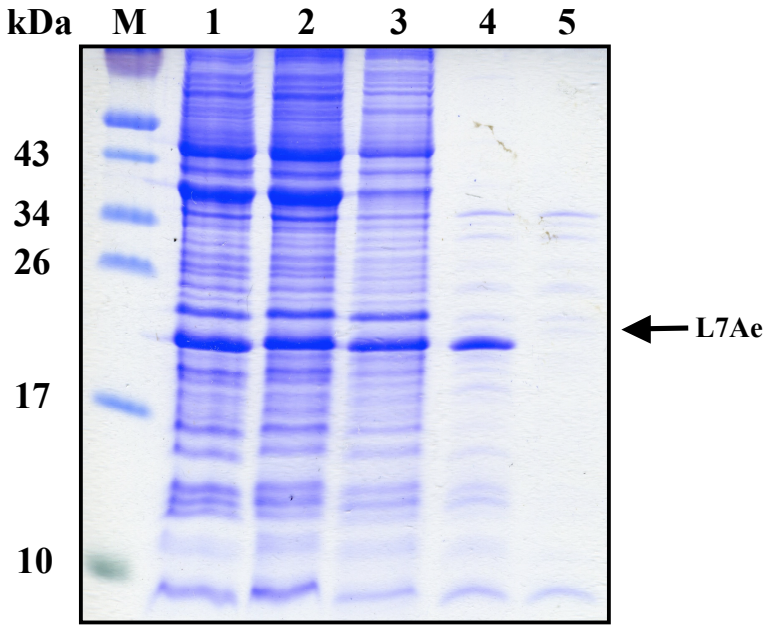


Figure 2.8. SDS-PAGE analysis of aliquots representative of the preliminary purification steps employed to isolate *Pae* L7Ae.

Pae L7Ae was identified in the soluble fraction after sonication (lane 1), PEI treatment (lane 2), and before/after heat treatment (lanes 3 and 4, respectively). *Pae* L7Ae was shown to bind a 1-ml HiTrap SP-Sepharose FF column due to its absence in the column flow-through (lane 5). *Pae* L7Ae was analyzed using 15% polyacrylamide tricine-SDS-PAGE and Coomassie blue staining. M: molecular weight markers/standards.

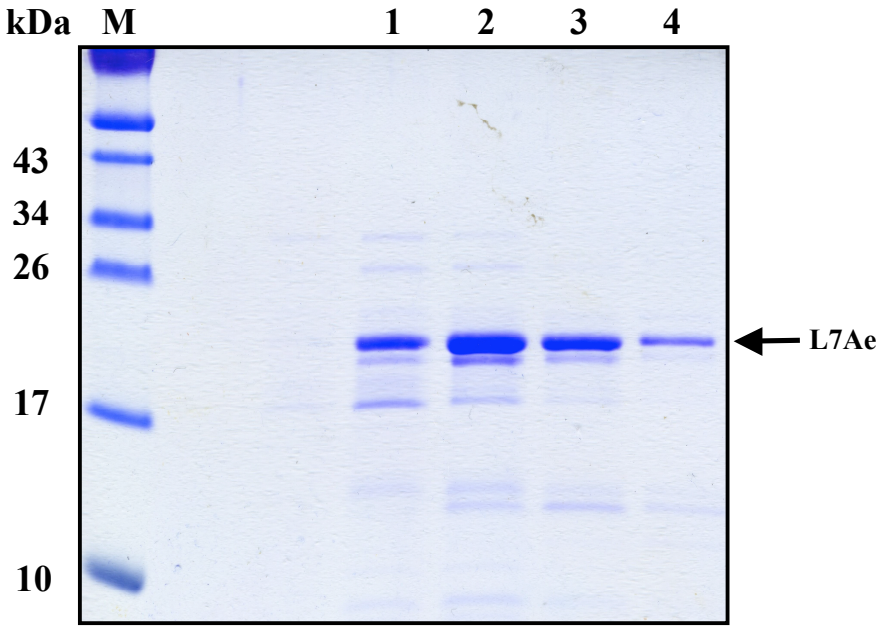


Figure 2.9. SDS-PAGE analysis of *Pae* L7Ae in the cation-exchange chromatography eluent fractions.

Pae L7Ae eluted from a 1-ml HiTrap SP-Sepharose FF column over a 0 to 2 M NaCl gradient. *Pae* L7Ae eluted in the range of approximately 0.9 - 1.3 M NaCl. Fractions containing *Pae* L7Ae were identified using SDS-PAGE and Coomassie blue staining. M: molecular weight markers/standards, 1-4: *Pae* L7Ae containing fractions.

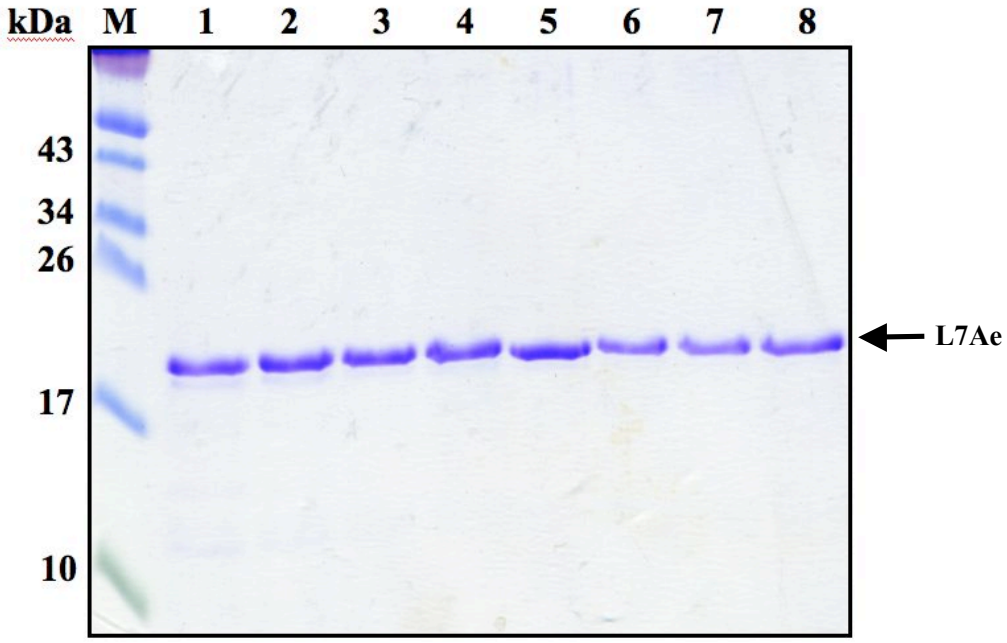


Figure 2.10. SDS-PAGE analysis of *Pae* L7Ae in the hydrophobic chromatography eluent fractions.

Pae L7Ae eluted from a 1-ml HiTrap phenyl-Sepharose HP column over a 2.5 to 0 M NaCl gradient. *Pae* L7Ae eluted at approximately 2.5 M NaCl. Fractions containing *Pae* L7Ae were identified using SDS-PAGE and Coomassie blue staining. M: molecular weight markers/standards, 1-8: *Pae* L7Ae containing fractions.

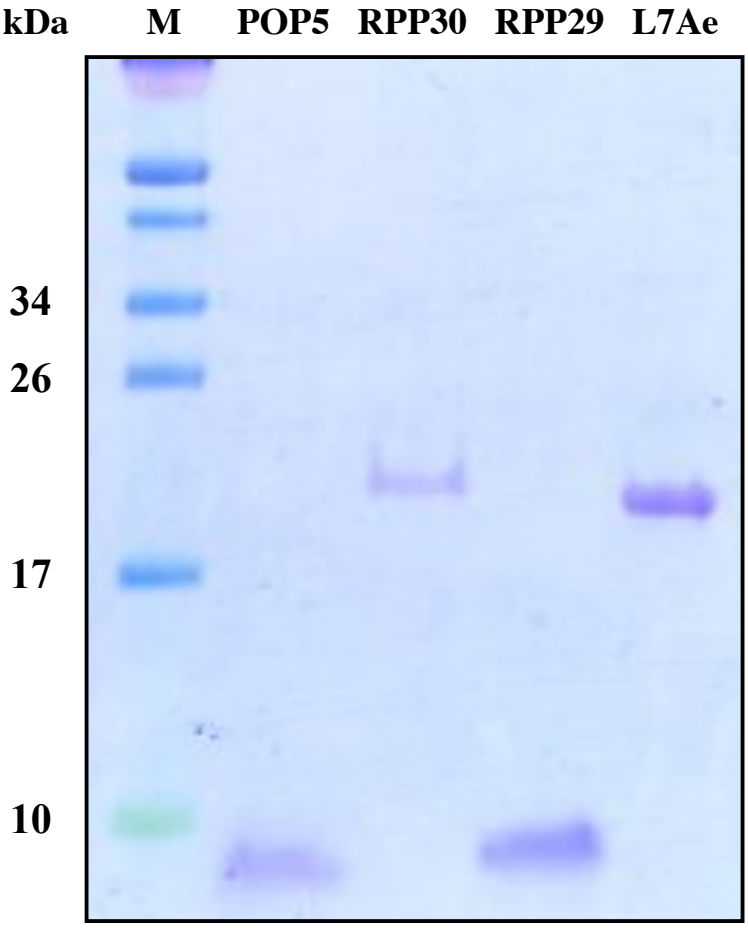


Figure 2.11. SDS-PAGE analysis of 2 - 4 µg of purified *Pae* RPPs.

Two to four µg of *Pae* POP5, RPP29, RPP30, and L7Ae were subjected to 15% (w/v) polyacrylamide SDS-PAGE and stained with Coomassie blue. M: molecular weight markers/standards

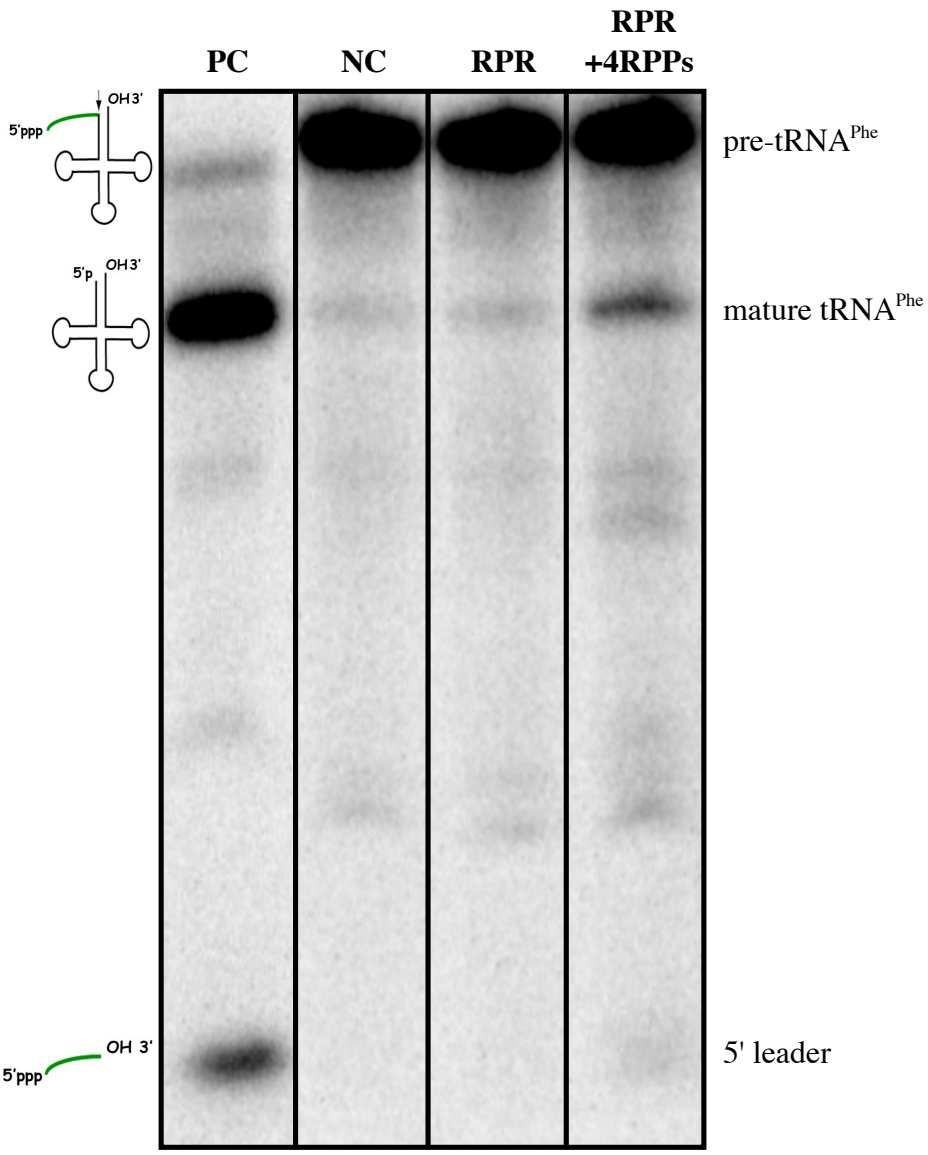


Figure 2.12. Functional reconstitution of the *Pae* RNase P holoenzyme.

Validation of *Pae* RNase P pre-tRNA processing activity by *in vitro* reconstitution using recombinant RNA and four purified subunits. PC (lane 1) is a positive control with pre-tRNA^{Phe} processed by *Eco* RNase P (95% cleavage). NC (lane 2) is a negative control in which pre-tRNA is incubated without RNase P (~2% cleavage/breakdown). RPR (lane 3) is an RNase P RNA-alone reaction (2% cleavage/breakdown). Although not shown, a reaction of the four *Pae* RPPs alone resembles the RPR alone control. RPR+4RPPs (lane 4) confirms pre-tRNA processing activity by *in vitro* reconstituted *Pae* RNase P holoenzyme (11% cleavage).

3. List of References

1. Altman, S. and Smith, J.D. (1971) Tyrosine tRNA precursor molecule polynucleotide sequence. *Nature*, **233**, 35-39.

2. Robertson, H.D., Altman, S. and Smith, J.D. (1972) Purification and properties of a specific Escherichia coli ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor. *J. Biol. Chem.*, **247**, 5243-5251.

3. Liu, F. and Altman, S. (2010) *Ribonuclease P. Protein Reviews Series.* New York, Springer-Verlag.

4. Gopalan, V. and Altman, S. (2006) Ribonuclease P: structure and catalysis. The RNA world (Gesteland, R.F., Cech, T.R. and Atlkins, J.F., Eds.), Cold Spring Harbor Laboratory Press, New York (http://rna.cshl.edu). Chapter 6.1.

5. Walker, S.C. and Engelke, D.R. (2006) Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit. Rev. Biochem. Mol. Biol.*, **41**, 77-102.

6. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, **35**, 849-857.

7. Schedl, P. and Primakoff, P. (1973) Mutants of Escherichia coli thermosensitive for the synthesis of transfer RNA. *Proc. Natl Acad. Sci. USA*, **70**, 2091-2095.

8. Ozeki, H., Sakano, H., Yamada, S., Ikemura, T. and Shimura, Y. (1975) Temperature-sensitive mutants of Escherichia coli defective in tRNA biosynthesis. *Brookhaven Symp. Biol.*, 89-105.

9. Tsai, H.Y., Masquida, B., Biswas, R., Westhof, E. and Gopalan, V. (2003) Molecular modeling of the three-dimensional structure of the bacterial RNase P holoenzyme. *J. Mol. Biol.*, **325**, 661-675.

10. Jovanovic, M., Sanchez, R., Altman, S. and Gopalan, V. (2002) Elucidation of structure-function relationships in the protein subunit of bacterial RNase P using a genetic complementation approach. *Nucleic Acids Res.*, **30**, 5065- 5073.

11. Reiter, N.J., Osterman, A., Torres-Larios A., Swinger, K.K., Pan, T. and Mondragon, A. (2010) Structure of bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*, **468**, 784-789.

12. Buck, A.H., Dalby, A.B., Poole, A.W., Kazantsev, A.V. and Pace, N.R. (2005) Protein activation of a ribozyme: the role of bacterial RNase P protein. *EMBO J.*, **24**, 3360-3368.

13. Crary, S.M., Niranjanakumari, S. and Fierke, C.A. (1998) The protein component of Bacillus subtilis ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNA^{Asp}. *Biochemistry*, **37**, 9409-9416.

14. Kurz, J.C. and Fierke, C.A. (2002) The affinity of magnesium binding sites in the Bacillus subtilis RNase P pre-tRNA complex is enhanced by the protein subunit. *Biochemistry*, **41**, 9545-9558.

15. Sun, L., Campbell, F.E., Zahler, N.H. and Harris, M.E. (2006) Evidence that substrate-specific effects of C5 protein lead to uniformity in binding and catalysis by RNase P. *EMBO J.*, **25**, 3998-4007.

16. Sun, L. and Harris, M.E. (2007) Evidence that binding of C5 protein to P RNA enhances ribozyme catalysis by influencing active site metal ion affinity. *RNA*, **13**, 1505-1515.

17. Guerrier-Takada, C. and Altman, S. (1992) Reconstitution of enzymatic activity from fragments of M1 RNA. *Proc. Natl Acad. Sci. USA*, **89**, 1266-1270.

18. Pan, T. (1995) Higher order folding and domain analysis of the ribozyme from Bacillus subtilis ribonuclease P. *Biochemistry*, **34**, 902-909.

19. Pomeranz-Krummel, D.A. and Altman, S. (1999) Verification of phylogenetic predictions in vivo and the importance of the tetraloop motif in a catalytic RNA. *Proc. Natl Acad. Sci. USA*, **96**, 11200-11205.

20. Loria, A. and Pan, T. (1997) Recognition of the T stem-loop of a pre-tRNA substrate by the ribozyme from Bacillus subtilis ribonuclease P. *Biochemistry*, **36**, 6317-6325.

21. Zahler, N.H., Christian, E.L. and Harris, M.E. (2003) Recognition of the 5' leader of pre-tRNA substrates by the active site of ribonuclease P. *RNA*, **9**, 734-745.

22. Gopalan, V. (2007) Uniformity amid diversity in RNase P. *Proc. Natl Acad. Sci. USA*, **104**, 2031-2032.

23. Hall, T.A. and Brown, J.W. (2002) Archaeal RNase P has multiple protein subunits homologous to eukaryotic nuclear RNase P proteins. *RNA*, **8**, 296-306.

24. Hartmann, E. and Hartmann, R.K. (2003) The enigma of ribonuclease P evolution. *Trends Genet.*, **19**, 561-569.

25. Ellis, J.C. and Brown, J.W. (2009) The RNase P family. *RNA Biol.,* **6**, 362-369.

26. Harris, J.K., Haas, E.S., Williams, D., Frank, D.N. and Brown, J.W. (2001) New insight into RNase P RNA structure from comparative analysis of the archaeal RNA. *RNA*, **7**, 220-232.

27. Pulukkunat, D.K. and Gopalan, V. (2008) Studies on Methanocaldococcus jannaschii RNase P reveal insights into the roles of RNA and protein cofactors in RNase P catalysis. *Nucleic Acids Res*., **36**, 4172-4180.

28. Kouzuma, Y., Mizoguchi, M., Takagi, H., Fukuhara, H., Tsukamoto, M., Numata, T. and Kimura, M. (2003) Reconstitution of archaeal ribonuclease P from RNA and four protein components. *Biochem. Biophys. Res. Commun.*, **306**, 666-673.

29. Tsai, H.Y., Pulukkunat, D.K., Woznick, W.K. and Gopalan, V. (2006) Functional reconstitution and characterization of Pyrococcus furiosus RNase P. *Proc. Natl Acad. Sci. USA*, **103**, 16147-16152.

30. Xu, Y., Amero, C.D., Pulukkunat, D.K., Gopalan, V. and Foster, M.P. (2009) Solution Structure of an Archaeal RNase P Binary Protein Complex: Formation of the 30-kDa Complex between *Pyrococcus furiosus* RPP21 and RPP29 Is Accompanied by Coupled Protein Folding and Highlights Critical Features for Protein–Protein and Protein–RNA Interactions. *J. Mol. Biol.*, **393**, 1043-1055.

31. Cho, I-M., Lai, L.B., Susanti, D., Mukhopadhyay, B. and Gopalan, V. (2010) Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc. Natl Acad. Sci. USA*, **107**, 14573-14578.

32. Fukuhara, H., Kifusa, M., Watanabe, M., Terada, A., Honda, T., Numata, T., Kakuta, Y. and Kimura, M. (2006) A fifth protein subunit Ph1496p elevates the optimum temperature for the ribonuclease P activity from *Pyrococcus horikoshii* OT3. *Biochem. Biophys. Res. Commun.,* **343**, 956-964.

33. Chen, W-Y., Pulukkunat, D.K., Cho, I-M., Tsai, H-Y. and Gopalan, V. (2010) Dissecting functional cooperation among protein subunits in archaeal RNase P, a catalytic ribonucleoprotein complex. *Nucleic Acids Res.*, **38**, 8316-8327.

34. Kawano, S., Nakashima, T., Kakuta, Y., Tanaka, I. and Kimura, M. (2006) Crystal structure of protein Ph1481p in complex with protein Ph1877p of archaeal RNase P from *Pyrococcus horikoshii* OT3: implication of dimer formation of the holoenzyme. *J. Mol. Biol.*, **357**, 583-591.

35. Honda, T., Kakuta, Y., Kimura, K., Saho, J. and Kimura, M. (2008) Structure of an archaeal homolog of the human protein complex Rpp21-Rpp29 that is a key core component for the assembly of active ribonuclease P. *J. Mol. Biol.,* **384**, 652-662.

36. Suryadi, J., Tran, E.J., Maxwell, E.S. and Brown, B.A. II (2005) The Crystal Structure of the *Methanocaldococcus jannaschii* Multifunctional L7Ae RNA-Binding Protein Reveals an Induced-Fit Interaction with the Box C/D RNAs. *Biochemistry*, **44**, 9657-9672.

37. Amero, C.D., Boomershine, W.P., Xu, Y. and Foster, M. (2008) Solution structure of Pyrococcus furiosus RPP21, a component of the archaeal RNase P holoenzyme, and interactions with its RPP29 protein partner. *Biochemistry*, **47**, 11704-11710.

38. Wilson, R.C., Bohlen, C.J., Foster, M.P. and Bell, C.E. (2006) Structure of Pfu Pop5, an archaeal RNase P protein. *Proc. Natl Acad. Sci. USA*, **103**, 873-878.

39. Hall, T.A. and Brown, J.W. (2004) Interactions between RNase P protein subunits in archaea. *Archaea*, **1**, 247-254.

40. Lai, L.B., Chan, P.P, Cozen, A.E., Bernick, D.L., Brown, J.W., Gopalan, V. and Lowe, T.M. (2010) Discovery of a minimal form of RNase P in Pyrobaculum. *Proc. Natl Acad. Sci. USA*, **107**, 22493-22498.

41. Lai, L.B., Cho, I-M., Chen, W-Y. and Gopalan, V. (2010). Archaeal RNase P: A mosaic of its bacterial and eukaryal relatives. Ribonuclease P: Protein Reviews Series (Liu, F. and Altman, S., Eds.). Springer-Verlag, New York, NY. **10**, 153-172.

42. Finkelstein, J., Antony, E., Hingorani, M.M. and O'Donnell, M. (2003) Overproduction and analysis of eukaryotic multiprotein complexes in Escherichia coli using a dual-vector strategy. *Anal. Biochem.*, **319**, 78-87.

43. Neff, M.M., Neff, J.D., Chory, J. and Pepper, A.E. (2002) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. *Plant J.,* **14,** 387-392.

44. Hardy, E. and Castellanos-Serra, L.R. (2004) "Reverse-staining" of biomolecules in electrophoresis gels: analytical and micropreparative applications. *Anal. Biochem.*, **328**, 1-13.

45. Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G. and Barral, Y. (2010) A role for codon order in translation dynamics. *Cell*, **141**, 355-367.